IIASA
International Institute for
Applied Systems Analysis
www.iiasa.ac.at

# Probability Theory in Geological Exploration

**Baecher, G.B.**

**IIASA Research Memorandum**
**May 1975**

RM-75-18

# PROBABILITY THEORY IN GEOLOGICAL EXPLORATION

Gregory Baecher

May 1975

# Probability Theory in Geological Exploration

Gregory Baecher

## Introduction

The purpose of this review is to summarize briefly certain applications of probability theory and statistics to geological exploration and inference, and in particular to problems of economic geology. The hope is that this summary will be of use in planning the IIASA conference on resource estimation scheduled for May, 1975, and will provide a reference within which to review certain contributions to that conference. For this purpose the discussion is organized under the following general headings:

1) General statistical applications
2) Spatial modelling
3) Regional mineral potential
4) Search theory in exploration
5) Subjective probability

## 1. General Statistical Applications

General statistical applications to geology have been made with increasing regularity since the time of Sir Ronald Fisher. Major contributions have been made by a relatively small but prolific group of geologists since the middle 1950's, among whom are Krumbein, Griffiths, and others. Much of this work is summarized in a series of basic texts on statistical methods in geology which include Krumbein and Graybill (1965), Miller and Kahn (1962), Griffiths (1967), and more recently Koch and Link (1971). In general, this work is directed more toward scientific inference than operations management or exploration strategies and is founded in frequentist probability theory.

Classical sampling theory, which appears in this literature, has been applied to problems of mineral exploration in a research program undertaken by the US Bureau of Mines and reported in the papers of Hazen, Becker, Berkenkotter, and others. This program has produced several reports on inference and sampling schemes for establishing ore grade and total ore content, but primarily represents a new application (i.e., a geological one) of already well-developed classical sampling theory rather than a unique analysis of exploration problems.

## 2. Spatial Modelling

A problem which has received considerable attention in the literature is the modelling of spatially distributed geological properties such as ore grade, magnetic potential, sediment composition, etc. Abstractly, let some physical property $Q(x,y,z)$ be associated with each point in the space $(x,y,z)$. The problem is to model this distribution of $Q$ with some set of mathematical relationships. Once such models are developed, they can be used to address problems of optimal sampling of the spatially distributed property and to predict the property at unobserved points.

It is often convenient to divide spatial distributions into large-scale trends and local randomness; this approach has led to what in the literature is called trend analysis (certainly, what constitutes a trend depends on the scale of interest). Trend analysis is in essence just a multivariate regression of Q on x, y and z, using common regression models of which the linear (i.e., simplest) case would be

$$Q(x,y,z) = b_o + b_1x + b_2y + b_3z + e \quad ,$$

where e is an error term called the "residual". In this case the trend component is a deterministic function with all randomness attributed to the residual term.

Trend analysis has been widely applied, and reviews are to be found in Krumbein and Graybill (1965), Harbaugh and Merriam (1968), and Agterberg (1974). Problems associated with applying trend analysis to geological data are similar to those of other regression problems, e.g., proper distribution families for residuals, heteroscedasticity, and autocorrelation of residuals; but as methods of multiple regression are well known, this methodology will not be further reviewed here. Applications to economic geology are found in Agterberg (1968), Forgotson (1963), Harbaugh (1964), and Hewlett (1964), among many others.

A second approach to spatial modelling is to assume that the property of interest can be represented by a stationary random function. Methods within this approach go under several names in the geological literature, but in essence are all extrapolations of Wienerian time-series analysis. In aggregate they might be referred to as stochastic spatial models, and they have counterparts in other disciplines treating spatial processes, e.g., geography and hydrology.

Stochastic spatial models assume stationarity, that is, they employ a random process model with mean, moments, and autocorrelation function constant over space. Matheron (1971, 1963) has developed models which only assume stationarity in differences of Q over space and thus are applicable to linearly varying means; however, real geological data usually exhibit both higher order deterministic variations and stochastic fluctuations, and only recently are combined models being developed (e.g., "Universal Kriging", Huijbregts and Matheron, 1970). These combined models typically proceed in three steps:

1. A deterministic trend surface is fitted by multiple regression.

2. The residuals from this trend surface are assumed to be manifestations of a stationary, zero-mean stochastic process, and the variance and autocorrelation functions are estimated.

3. Estimates of aggregate properties or predictions are made on the residuals, and the trend component is reintroduced.

Theoretical problems are associated with this procedure (Watson, 1971, Agterberg, 1974), but at present it is a substantial although pragmatic tool for treating spatial problems.

Trend analysis and stochastic spatial models have been used primarily for drawing inferences from sample data rather than for optimizing sampling plans or predicting values at unobserved locations; however, some applications to predicting sparsely sampled regions within intensively sampled areas have been made by Krige (1966). The main application thus far has been estimating integrals of the spatial surface (e.g., total ore deposit value).

Stochastic spatial process models have been used in fields outside geology to optimize sampling strategies, among which is the classic work of Matern (1960) in forestry and the recent extensions of his work to rainfall measurement by Rodriquez-Iturbe and Mejia (1973).

One is hard pressed to see how stochastic spatial models could be applied to "resource" (as opposed to "reserve") estimates, as they are based on the concept of a homogeneous, continuous function of the variable. Of course many formations of interest (e.g., coal beds) extend over large areas, and modelling their properties with spatial process models may yield better estimates of their total content. Work has also been done on associating residual maps from trend analysis with mineralization, there being some evidence of a correlation between areas of high autocorrelation in residuals and ore bodies.

3. Regional Mineral Potential

Of great interest for the question of resource estimation is the body of work on estimating so-called "regional mineral potential." Resource potential is usually defined as some monetary measure of the mineral deposits which are predicted to exist within given regions; sometimes this is broken down into particular resource types, but often potential is an aggregate through the common index of monetary value.

A straightforward way of making potential estimates is that initially undertaken by Harris (1965), in which a multivariate regression of regional mineral value is made on several geological variables using an extensively explored region as control. That regression is then used predictively to estimate mineral potential for non-intensively explored regions of approximately similar geology. Problems of using this empirical approach and of extending findings to new areas are clear. Singer (1972) made the interesting finding in his application of factor analysis that regional economic value based on known reserves is most heavily dependent on intensity of development and is almost independent of geology. Other applications of this approach have been made by Agterberg (1971), Agterberg, et al. (1972), Sinclair and Woodsworth (1970), and DeGeoffroy and Wignall (1970).

Similar to the regression, factor analysis approach, but simpler, are methods in which the density of mineral deposits in known portions of geological formations is extended to unexplored portions; these have been developed by DeGeoffroy and Wu (1970) who computed the known valuation of mineral deposits in certain units of the Canadian Shield and extended this valuation to undeveloped portions. Several applications similar in intent to this have been presented in the Stanford Conference (Crandall and Harbaugh, 1974) for making oil resource estimates.

A totally different approach from the regression, factor analysis one is that of Allais (1957) and Kaufman (1963). There is substantial evidence to suggest that a log-normal distribution is widely applicable to geological data. Among other things, the size distribution of mineral deposits and particularly oil reservoirs is well modelled by this family of distributions. There have been several confirmations of the latter conclusion (Kaufman, 1963, 1965; Uhler and Bradley, 1970; Arps and Roberts, 1958). If one assumes that deposits are randomly located in two-dimensional space, the number of occurrences within elemental units of equal size can be modelled by the Poisson distribution. Therefore, if estimates of distribution parameters can be made on the basis of currently known reserves, an estimate of the total resource within a geological unit can be made. This approach has been used by several authors and results in the most rigorously derived estimates of resources (Kaufman, 1974). Assumption of these distribution types can also serve as the basis for developing optimal strategies of search for new deposits (e.g., Allais, 1957).

Work by Griffiths (1966) and DeGeoffroy and Wu (1970) seems to indicate that a negative binomial distribution for location may be more appropriate than a Poisson distribution, as deposits are not random but tend to cluster; however, this does not change the methodological approach.

4. Search Theory in Mineral Exploration

Stochastic process models are primarily aimed at mapping and pattern recognition of geological variables. A second problem of even greater concern in exploration is the search for as yet undiscovered geological bodies or mineral resources. The operational problem is how to optimally allocate exploration effort over time, space, and exploration tools either to maximize the probability of finding a new body or to minimize the degree of uncertainty about the location of undiscovered bodies (these two objectives do not necessarily lead to identical optimal allocations).

In the English literature, rigorous approaches to spatial search began to appear with the publication of the U. S. Office of Naval Research study under the direction of B. O. Koopman (1956a, 1956b, 1957). Koopman search, as most search theories, is based on Bayesian probability theory. In this theory, a prior probability density function of target location is generated either on the basis of subjective information, historical data, or mineral potential studies, and then given a detection function for the search tool being used. A continuous spatial allocation is derived which maximizes the integral

$$P(\text{find}) = \int_{x_o} \text{pdf}(x_o)\ P\left(\text{find}/x_o, \varphi_{x_o}\right) dx_o$$

in which $\text{pdf}(x_o)$ is the probability density function of target location, and $P\left(\text{find}/x_o, \varphi_{x_o}\right)$, the detection function, is the conditional probability of finding a target at location $x_o$ given that $\varphi_{x_o}$ amount of exploration effort is allocated to $x_o$ and that a target exists there. In Koopman search, the detection function is assumed to have the form

$$P\left(\text{find}/x_o, \varphi_{x_o}\right) \propto \left[1 - \exp\left(-\varphi_{x_o}\right)\right] .$$

DeGuenin (1961) extended this theory to any detection function having diminishing returns.

Search theory has been applied to a variety of problems from police reconnaissance to library browsing, as well as to geological exploration (DeGuenin, 1961, Brown, 1960). Morse (1974) has recently written a state-of-the-art of search theory.

In application, one may use comparatively simple graphical procedures for optimizing search allocations with both the Koopman theory and DeGuenin's extension. However, as far as I know, nomographs and charts which might make this procedure even easier have yet to appear, unless they exist as proprietary information within exploration companies. This means that optimization must be done on a case-by-case basis.

A special case of the search allocation problem which finds application in geological exploration is that of using a relatively inexpensive search tool (e.g., geophysical surveys) as a first stage before committing expensive tools (e.g., borings). The first stage serves to "concentrate" the prior probability distribution so that second stage effort may be allocated more judiciously. Classic applications of this are Allais' (1957) work on exploration of the Sahara, and Engle's (1957) and Slichter's (1955) work on geophysics in exploration. Stone et al. (1972) have developed mathematical aspects of the false target problem in two stage search.

A subclass of search problem which has received more attention than most is that of drilling patterns. Grid drilling allocates exploration effort uniformly over the region searched; it grades into optimal search when the region is decomposed into elements, and grid spacing within each element is determined on the a priori probability (or mineral potential) of a deposit existing in that element (i.e., more probable elements have smaller grid spacings). Charts and nomograms for grid drilling for a number of target shapes and grid geometries are to be found in

Singer and Wickman (1969), Savinskii (1965), and Slichter (1955). Drew (1966) has analyzed the proposal of large-scale (i.e., sub-continental) grid drilling for national mineral resources.

An interesting problem arising out of search theory and estimating mineral potential is the contact between exploration and economic policy; hence the prospect both of making large-scale national investments in exploration and of optimally allocating that investment within exploration. A typical example is ocean floor exploration: How much are we justified in investing in exploration of the ocean floor in order to improve economic decision-making, and how do we allocate it? Clearly the decision depends on what we might expect to learn from investments of given sizes, which in turn can be estimated from search theory and mineral potential studies. An approach in this direction has been made by MacAvoy and Pindyck (1974) who have included Kaufman's predictions of the returns from large-scale exploration strategies in their econometric modelling.

## 5. Subjective Probability

Geological exploration is fundamentally an inductive process which depends heavily on experience, intuition, and familiarity with geological processes--despite what the literature of "geomathematics" would have one believe. In my opinion the models and operational procedures applied to this process serve only to combine the "feelings" of geologists logically with subsequent empirical data in ways which lead to justifiable and sound inferences; and to indicate how, by starting from those "feelings", future effort might be optimally allocated. If this is the case, then the theories of subjective probability and Bayesian induction ought to have major applicability to exploration. However, the geomathematical literature has been remarkably slow, given the state of subjectivist theory, to make use of the so-called Bayesian approach and the resulting statistical decision theory.

Grayson's (1960) early and well-known application of subjectivist theory to wild-catting decisions in the oil industry clearly indicate the applicability of statistical decision analysis not only to economic geology, but to industrial decisions generally. Kaufman (1962) extended this approach and refined the mathematical basis of estimation and decision on oil exploration. But beyond these initial attempts the Bayesian approach has seen less effort allocated to it than, again in my opinion, less promising classical statistical applications. With subsequent development, of both subjectivist theory (and practice) and utility theory, we now have a refined set of methodologies with which to look again at subjectivist techniques in exploration. These methods are further strengthened by intervening work in the reliability and accuracy of subjective assessments by Edwards (1954) and many others.

Although there have been some shoddy attempts to apply subjectivist theory to exploration, two recent applications may lead to further work. The first is that of Folayan et al. (1970) in engineering exploration of soil deposits, and the second that of Harris et al. (1971) and Barry and Freyman (1970) on mineral estimates of the Canadian Shield. Both are straightforward uncomplicated applications to estimating geological conditions on the basis of "expert" opinion and both indicate promising avenues for further work.

References

[1] Agterberg, F.D., Geomathematics. Elsevier (1974).

[2] Agterberg, F.D., "A probability index for detecting favourable geological environments." Can. Inst. Min. Metall., Spec. Vol. 12: 82-91, (1971).

[3] Agterberg, F.D., "Application of trend analysis in the evaluation of the Whalesback Mine, Newfoundland." Can. Inst. Min. Metall., Spec. Vol. 9: 77-88, (1968).

[4] Agterberg, F.D., et al., "Geomathematical evaluation of copper and zinc potential of the Abitibi area, Ontario and Quebec." Geol. Surv. Can., Paper 71-41: p. 55, (1972).

[5] Allais, M., "Method of appraising economic prospects of mining exploration over large territories; Algerian Sahara case study." Mgt. Sci., 3: 285-347, (1957).

[6] Arps, J.J. and Roberts, T.G., "Economics for drilling for cretaceous oil production on the east flank of the Denver-Julesberg Basin." Bull. AAPG, 42 (11), (1958).

[7] Barry, G.S. and Freyman, A.J., "Mineral endowment of the Canadian Northwest--A subjective probability assessment." Can. Min. Metall. Bull., 63: 1031-1042, (1970).

[8] Brown, A.A., "Search theory and problems of exploration drilling." Bull. Min. Ind. Exp. Sta., Penn. State Univ. 72: 33-37, (1960).

[9] Crandall, K.H. and Harbaugh, J.W. (eds.), Oil Resource Estimation. Stanford University, Dept. of Applied Earth Sciences, (1974).

[10] DeGeoffroy, J.G. and Wignall, T.K., "A probabilistic appraisal of mineral resources in a portion of the Grenville Province of the Canadian Shield." Econ. Geol., 66: 466-479, (1970).

[11] DeGeoffroy, J.G. and Wu, S.M., "A statistical study of ore occurrences in the greenstone belts of the Canadian Shield." Econ. Geol., 65: 496-509, (1970).

[12] DeGuenin, J., "Optimum distribution of effort: an extension of the basic Koopman theory." Oper. Res. 9: 1-7, (1961).

[13] Drew, L.J., Grid Drilling Exploration and its Application to the Search for Petroleum. PhD Dissertation, Penn. State Univ., (1966).

[14] Edwards, W., "The theory of decision making." Psychological Bull., 51: 380-417, (1954).

[15] Engle, J.H., "Use of clustering in mineralogical and other surveys." 1st Int'l Conf. on Oper. Res., Oxford, (1957).

[16] Folayan, J.I., et al., "Decision theory applied to settlement predictions." J. Soil Mech. and Foundations, ASCE SM4: 1127-41, (1970).

[17] Forgotson, J.M., "How computers help find oil." Oil Gas J., 61: 100-109, (1963).

[18] Grayson, C.J., Decisions under Uncertainty: Drilling Decisions by Oil and Gas Operators. Harvard U. Pr., p. 402, (1960).

[19] Griffiths, J.C., Scientific Method in the Analysis of Sediments. McGraw-Hill, pg. 508, (1967).

[20] Griffiths, J.C., "Exploration for natural resources." Oper. Res., 14: 189-209, (1966).

[21] Harbaugh, J.W., "Trend surface mapping of hydrodynamic oil traps with the IBM 7090-94 computer." Colo. Sch. Min. Qt., 59: 557-78, (1964).

[22] Harbaugh, J.W. and Merriam, D.F., Computer Applications in Stratiagraphic Analysis. Wiley, p. 282, (1968).

[23] Harris, D.P. and Euresty, D.A., "The impact of transportation network upon the potential supply of base and precious metals." Proc. 10th. Int. Symp. Appl. Comput. Methods Miner. Ind., S. Afr. Institute Min. Metall: 99-108, (1973).

[24] Harris, D.P., et al., "A mineral resource appraisal of the Canadian Northwest using subjective probabilities and geological opinion." Can. Inst. Min. Metall., Spec. Vol., 12: 100-116, (1971).

[25] Harris, D.P., An Application of Multivariate Statistical Analysis to Mineral Exploration. PhD Dissertation, Penn. State Univ., (1965).

[26] Hewlett, R.F., "Polynomial surface fitting using sample data from an underground copper deposit." USBM Rept. Inv. 6522, p.27, (1964).

[27] Huijbregts, C. and Matheron, G., "Universal Kriging: an optimal method for estimating and contouring in trend surface analysis." 9th. Int'l Symp. on Tech. for Dec. Making in Min. Ind., (1970).

[28] Kaufman, G., Statistical Decision and Related Techniques in Oil and Gas Exploration. Prentice-Hall, p. 307, (1963).

[29] Koch, G.S. and Link, R.F., Statistical Analysis of Geological Data. Wiley, 2 vols., (1971).

[30] Koopman, B.O., "The theory of search." Oper. Res., 4: 324-346 (I), 4: 503-531 (II), 5: 613-626 (III) (1956a, b; 1957).

[31] Krige, D.B., "A study of gold and uranium distribution patterns in the Klerkdorp goldfield." Geo-exploration, 4: 43-53, (1966).

[32] Krumbein, W.C. and Graybill, R.A., An Introduction to Statistical Models in Geology. McGraw-Hill, p. 475, (1965).

[33] MacAvoy, P. and Pindyck, R.S., "Alternative regulatory policies for dealing with the natural gas shortage." Bell. J. Mgt. Sci., (1974).

[34] Matern, B., Spatial Variation, Medd. Skogsforskn Inst. 49 (5): p. 144, (1960).

[35] Matheron, G., "The theory of regionalized variables and its application." Cah. Cent. Morphol. Math., 5: p. 211, (1971).

[36] Matheron, G., "Principles of Geostatistics." Econ. Geol., 58: 1246-56, (1963).

[37] Miller, R.L. and Kahn, J.S., Statistical Analysis in the Geological Sciences. Wiley, p. 483, (1962).

[38] Morse, P.M., Search Theory. MIT Oper. Res. Center working paper, 74-1, (1974).

[39] Rodriquez-Iturbe, I. and Mejia, J.M., Design of Rainfall Networks in Time and Space. Preprint of a paper to appear in Water Resources Research, (1973).

[40] Romanova, M.A. and Sarfmnov, O.V. (eds.), Topics in Mathematical Geol. Consultants Bureau, N.Y., (1970).

[41] Savinskii, I.D., Probability Tables for Locating Elliptical Underground Masses with a Rectangular Grid. Consultants Bureau, N.Y., (1965).

[42] Serra, J., "Morphologie mathematique et interpretation metallogenique." Mineral. Deposita, 3: 135-54, (1968).

[43] Sinclair, A.J. and Woodsworth, G.L., "Multiple regression as a method of estimating exploration potential in an area near Terrace, B.C." Econ. Geol., 65: 998-1003, (1970).

[44] Singer, D.A., Multivariate Statistical Analysis of the Unit Regional Value of Mineral Resources. Thesis, Penn. State Univ., (1972).

[45] Singer, D.A. and Wickman, F.E., "Probability tables for locating elliptical targets with square, rectangular and hexagonal point nets." Min. Sci. Exp. Sta., Penn. State Univ. special publication 1-69, (1969).

[46] Slichter, L.B., "Geophysics applied to prospecting for ores." Econ. Geol., 50th Anniv. Vol: 885-969, (1955).

[47] Stone, L.D., et al., "Optimal search in the presence of Poisson-distributed false targets." SIAM J. Appl. Math., 23: 6-27, (1972).

[48] Uhler, R.S. and Bradley, P.G., "A stochastic model for petroleum exploration over large regions." J. Am. Stat. Assoc., 65: 623-630, (1970).

[49] Vistelius, A.B., Mathematical Geology. Systematic Index of Reference. Lab. Math. Geol., Leningrad, p. 246, (1969).

[50] Vistelius, A.B., "On checking of the theoretical backgrounds of the stochastical models in concrete geological investigations." Int. Geol. Congr., 23rd, Prague, Proc. 13: 153-161, (1968).

[51] Watson, G.S., "Trend surface analysis." Int'l J. Assn. Math. Geol., 3: 215-226, (1971).